

Ejemplo 1:

Se están estudiando tres procesos (A, B, C) para fabricar pilas o baterías. Se sospecha que el proceso incide en la duración (en semanas) de las baterías, es decir, que la duración (en semanas) de los procesos es diferente. Se seleccionan aleatoriamente cinco baterías de cada proceso y al medirles aleatoriamente su duración los datos que se obtienen son los siguientes:

- 1) Activa tu directorio de trabajo.
- 2) Crea un nuevo script y llámale "Script10-DatosBivariados2"
- 3) Crea un vector de datos para cada proceso descrito en el problema.

```
A <- c(100,96,92,96,92)
B <- c(76,80,75,84,82)
C <- c(108,100,96,98,100)

A;B;C

## [1] 100 96 92 96 92
## [1] 76 80 75 84 82
## [1] 108 100 96 98 100
```

4) Crea una hoja de datos teniendo como componentes (columnas) los tres vectores (se puede hacer pues el número de datos en cada proceso es igual, de lo contrario se debería de crear dos variables una para la duración de cada proceso y otra para identificar a qué proceso corresponde).

```
Baterias <- data.frame(procesoA=A, procesoB=B, procesoC=C)
Baterias

##   procesoA procesoB procesoC
## 1     100      76     108
## 2      96      80     100
## 3      92      75      96
## 4      96      84      98
## 5      92      82     100

# Para editar los datos puede utilizar la función fix()
fix(baterias)
```

- 5) Guarda la hoja de datos en un archivo.

```
write.table(Baterias, file="Baterias.txt", append=FALSE, quote=TRUE, sep=" ",
            na="NA", col.names=TRUE)
```

- 6) Elimina todos objetos que existen en el espacio de trabajo (Workspace)

```
ls()

## [1] "A"          "B"          "baterias" "Baterias" "C"

rm(list=ls(all=TRUE))
ls()

## character(0)
```

7) Recupera la hoja de datos, para probar si fue guardada.

```
Baterias <- read.table("Baterias.txt", header=TRUE)
Baterias

##   procesoA procesoB procesoC
## 1      100       76      108
## 2       96       80      100
## 3       92       75       96
## 4       96       84       98
## 5       92       82      100
```

8) Conecta o adjunta la hoja de datos a la segunda ruta o lista de búsqueda.

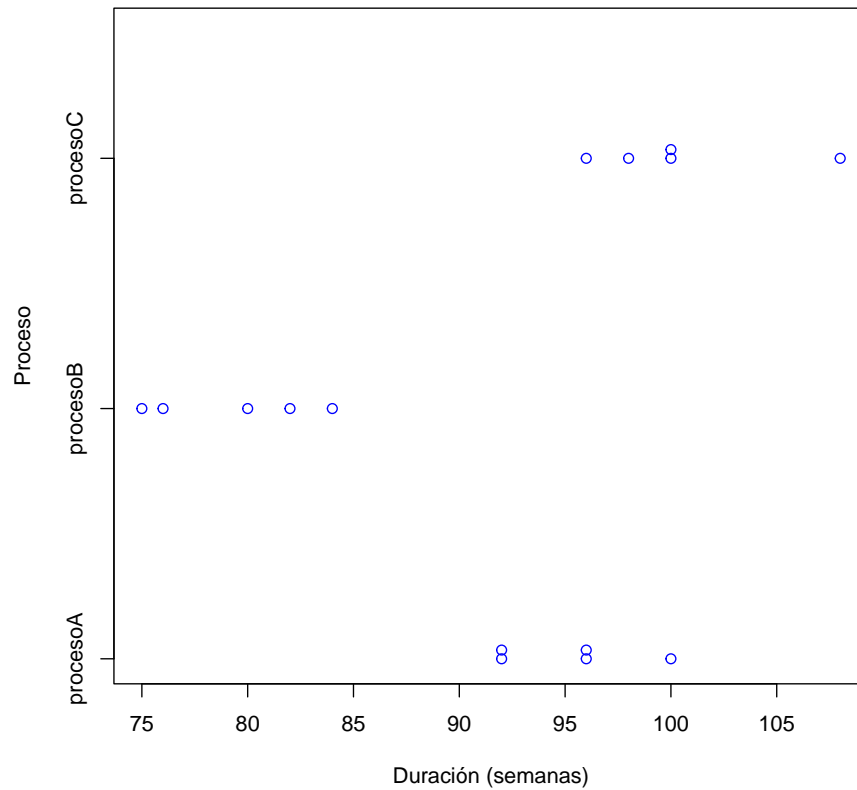
```
attach(Baterias, pos=2)
search()

## [1] ".GlobalEnv"      "Baterias"         "package:knitr"
## [4] "package:stats"    "package:graphics" "package:grDevices"
## [7] "package:utils"    "package:datasets" "package:methods"
## [10] "Autoloads"        "package:base"
```

9) Dibuja un gráfico horizontal de puntos para los tres procesos.

```
stripchart(Baterias, main="Gráfico de puntos para los tres procesos", method = "stack",
           vertical =FALSE, col="blue", pch=1, xlab="Duración (semanas)",
           ylab="Proceso")
```

Gráfico de puntos para los tres procesos



*# Note que con ayuda de este gráfico podemos observar si los tres procesos se comportan
#de manera distinta o parecida en cuanto a duración en semanas de las baterías.*

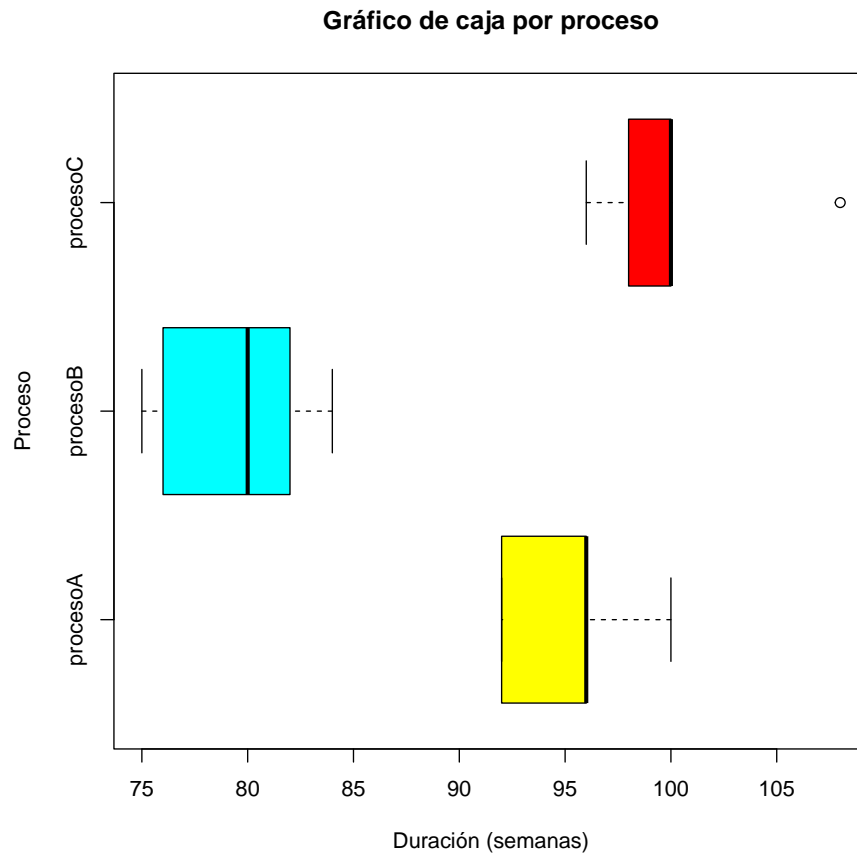
10) Muestra un resumen estadístico para los tres procesos.

```
summary(Baterias)
```

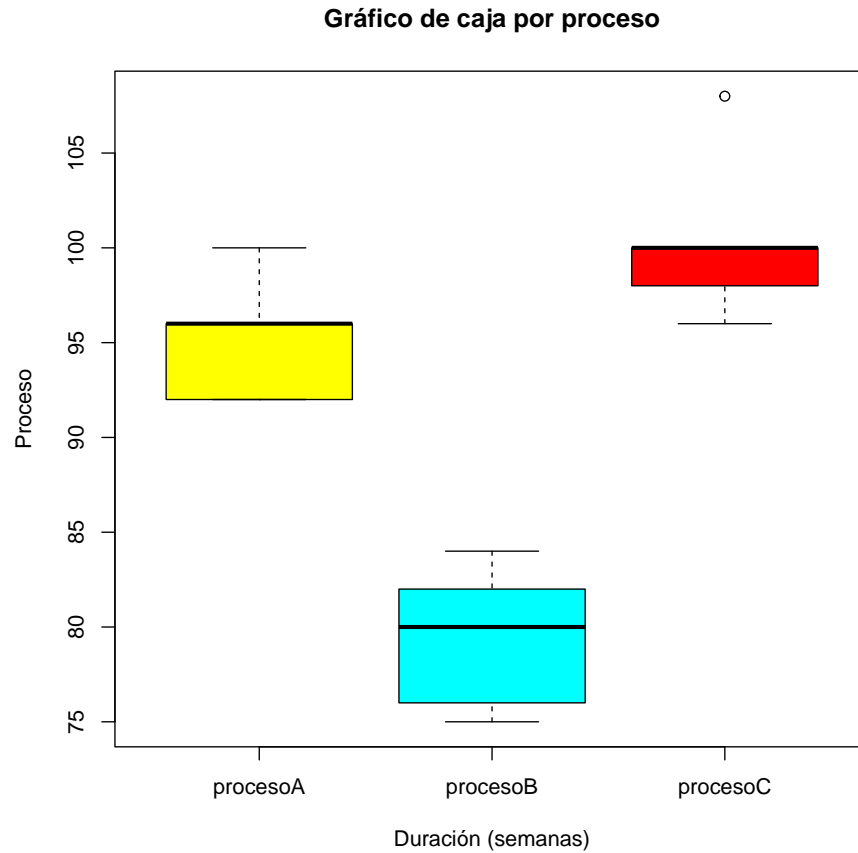
##	procesoA	procesoB	procesoC
## Min.	: 92.0	Min. :75.0	Min. : 96.0
## 1st Qu.:	92.0	1st Qu.:76.0	1st Qu.: 98.0
## Median :	96.0	Median :80.0	Median :100.0
## Mean :	95.2	Mean :79.4	Mean :100.4
## 3rd Qu.:	96.0	3rd Qu.:82.0	3rd Qu.:100.0
## Max. :	100.0	Max. :84.0	Max. :108.0

11) Dibuja un gráfico de cajas (box-plot) para los tres procesos.

```
# Horizontal
boxplot(Baterias, width=NULL, varwidth=TRUE, names, add= FALSE, horizontal = TRUE,
        main="Gráfico de caja por proceso", border=par("fg"),
        col=c("yellow", "cyan", "red"),
        xlab = "Duración (semanas)", ylab="Proceso")
```



```
# Vertical
boxplot(Baterias, width=NULL, varwidth=TRUE, names, add= FALSE, horizontal = FALSE,
        main="Gráfico de caja por proceso", border=par("fg"),
        col=c("yellow", "cyan", "red"),
        xlab = "Duración (semanas)", ylab="Proceso")
```



12) Presenta la matriz de covarianzas muestral.

```
options(digits=3) # sólo imprime 3 lugares decimales
S <- var(Baterias)
S

##           procesoA procesoB procesoC
## procesoA      11.2      -1.6      12.4
## procesoB      -1.6      14.8      -4.7
## procesoC      12.4      -4.7      20.8
```

13) Presenta la desviación estándar de cada proceso.

```
desv <- sd(Baterias)

## Error in is.data.frame(x): 'list' object cannot be coerced to type
## 'double'
```

```
desv
```

```
## Error in eval(expr, envir, enclos): objeto 'desv' no encontrado
```

14) Realiza un análisis de varianza de una vía, para probar la hipótesis nula de que el proceso no influye en la duración de las baterías, es decir, que no hay diferencias entre los tres procesos.

$H_0 : \mu_A = \mu_B = \mu_C$, no existe diferencias entre los tres procesos.

$H_a : \mu_i \neq \mu_j$, por lo menos un par $i \neq j$, de procesos difieren en la duración de las baterías.

```
# Concatena los tres vectores dentro de un vector simple, junto con un vector factor  
# indicador de la categoría o tratamiento (A, B, C) que origina cada observación.  
# El resultado es un data.frame que tiene como componentes los dos vectores anteriores.
```

```
Baterias <- stack(Baterias)
```

```
Baterias
```

```
##      values      ind  
## 1      100 procesoA  
## 2       96 procesoA  
## 3       92 procesoA  
## 4       96 procesoA  
## 5       92 procesoA  
## 6       76 procesoB  
## 7       80 procesoB  
## 8       75 procesoB  
## 9       84 procesoB  
## 10      82 procesoB  
## 11      108 procesoC  
## 12      100 procesoC  
## 13       96 procesoC  
## 14       98 procesoC  
## 15      100 procesoC
```

```
names(Baterias) # Muestra los encabezados de los vectores
```

```
## [1] "values" "ind"
```

```
# Prueba de igualdad de medias por descomposición de la varianza en dos fuentes de variación  
# la variabilidad que hay entre los grupos (debida a la variable independiente o los  
# tratamientos), y la variabilidad que existe dentro de cada grupo (variabilidad no  
# explicada por los tratamientos).
```

```
aov.Baterias <- aov(values~ind, data=Baterias)
```

```
# values~ind relaciona los valores muestrales con los respectivos grupos
```

```
summary(aov.Baterias)

##              Df Sum Sq Mean Sq F value    Pr(>F)
## ind              2   1196      598   38.3 6.1e-06 ***
## Residuals      12    187       16
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# Note que es necesario la instrucción anterior para poder visualizar la tabla ANOVA

# Prueba de igualdad de medias en un diseño de una vía (o unifactorial) asumiendo que las
# varianzas de los grupos son iguales
oneway.test(values~ind, data=Baterias, var.equal = TRUE)

##
## One-way analysis of means
##
## data:  values and ind
## F = 38, num df = 2, denom df = 12, p-value = 6e-06
```

Decisión: ya que $\alpha = 0.05 > p\text{-value}$ obtenido, entonces se rechaza H_0

15) Deshace la concatenación del vector de valores y el vector indicador de categoría.

```
Baterias = unstack(Baterias)
Baterias

##   procesoA procesoB procesoC
## 1      100       76      108
## 2       96       80      100
## 3       92       75       96
## 4       96       84       98
## 5       92       82      100
```

16) Desconecta la hoja de datos de la segunda ruta o lista de búsqueda.

```
detach(Baterias, pos=2)
search()

## [1] ".GlobalEnv"      "package:knitr"    "package:stats"
## [4] "package:graphics" "package:grDevices" "package:utils"
## [7] "package:datasets" "package:methods"  "Autoloads"
## [10] "package:base"
```